Ramesh Bagadi Ideas Journal
Volume 1, Issue 5
(October 2021)

Applicability Criterion For K-Means Clustering Algorithm

RAMESH CHANDRA BAGADI

ISBN: 9798542222288

DEDICATION

This text is dedicated to the all compassionate *Creator* of the *Universe*.

CONTENTS

ACKNOWLEDGMENTS

The author would like to express his deepest gratitude to all the members of his Loving Family, especially his ever loving son Rohith Rishi Bagadi, Respectable Teachers, En-Dear-Able Friends, Inspiring Social Figures, Highly Esteemed Professors, Reverence Deserving Deities that have deeply contributed in the formation of the necessary scientific temperament and the social and personal outlook of the author that has resulted in the conception, preparation and authoring of this research manuscript note. The author pays his sincere tribute to all those dedicated and sincere folk of academia, industry and elsewhere who have sacrificed a lot of their structured leisure time and have painstakingly authored treatises on Science, Engineering, Mathematics, Art and Philosophy covering all the developments from time immemorial until then, in their supreme works. It is standing on such treasure of foundation of knowledge, aided with an iota of personal god-gifted creativity that the author bases his foray of wild excursions into the understanding of natural phenomenon and forms new premises and scientifically surmises plausible laws. The author strongly reiterates his sense of gratitude and infinite indebtedness to all such '*Philosophical Statesmen*' that are evergreen personal librarians of Science, Art, Mathematics and Philosophy.

1 INTRODUCTION

In this research investigation, firstly the author presents a Criterion Of Applicability Of K-Means Clustering Algorithm On A Given Data Based On The Limit Of Variation Of Results Of Several Runs Of K-Means Clustering Algorithm. It should be noted that the results are data specific and can change from data to data.

2 THE K-MEANS CLUSTERING ALGORITHM

2.1 The K-Means Clustering Algorithm

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume K clusters) fixed apriori. The main idea is to define K centers, one for each cluster. These centers should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate K new centroids as barycenter of the clusters resulting from the previous step. After we have these K new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop we may notice that the K centers change their location step by step until no more changes are done or in other words centers do not move any more. Finally, this algorithm aims at minimizing an objective function known as squared error function.

There is a lot of literature on K-Means Clustering available on K-Means Clustering. Some good sources are [1], [2],[3], and [4]. Also Cluster Validation Measures such as SSW (Sum of Squares Within), Elbow Plot and Silhouette Plot are detailed well in these afore-referred sources.

For a given dataset of points, K-Means Clustering aims at finding

clusters in the data.

The clusters are found using the following procedure:

For any clustering problem, the number of clusters is to be a given quantity. Therefore, let us assume that we need to find K number of clusters from among the *n* data points.

Step 1: Firstly, we randomly pick K number of random centers among the *n* data points that are to act as the K clusters centroids.

Step 2: Now, for each such cluster centroid, we assign points to it that are nearest to this cluster centroid than any other cluster centroid.

Step 3: We now compute the new centroids of points belonging to each cluster again after such aforementioned assignments.

Step 4: We now repeat the algorithm from Step 2 onwards and keep repeating this procedure till

a) the Cluster centroids do not change anymore, i.e., they converge to some values.

b) the points of a cluster remain in the same cluster.

c) maximum number of iterations are reached. This number is pre-selected at the outset of the beginning of the algorithm.

2.2 The K-Means Clustering Objective Function

The objective of K-Means clustering is to minimize total intra-cluster variance, i.e., the squared error function:

$$J = SSE = \sum_{j=1}^{K} \sum_{i=1}^{n_j} \left\| {}^{j}x_i - c_j \right\|^2$$

where

${}^{j}x_i$ is the i^{th} point of the j^{th} cluster, c_j is the centroid of the j^{th} cluster, K is the number of clusters and n_j is the number of elements of the j^{th} cluster.

Also, c_j is given by

$$c_j = \left\{ \frac{\sum_{i=1}^{n_j} \left({}^{j}x_i \right)}{n_j} \right\}$$

2.3 Cluster Evaluation

2.3.1 Compactness

Sum of squares within clusters (SSW) or within cluster variance is given by

$$SSW = \sum_{i=1}^{n_j} \left\{ \left\| {}^{j}x_i - c_j \right\|^2 \right\}$$

The index can only be used for numerical data because it requires centroids of clusters. SSW measures the compactness of clusters, and is suitable for centroid-based clustering, where hyperspherical clusters are desired. The value of SSW always decreases as the number of clusters

increases.

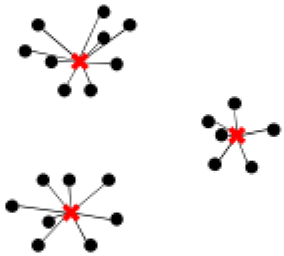


Fig: *Illustration of Sum Of Squares Within Clusters*

2.3.2 Dunn's Index

Many clustering algorithms require the number of clusters given as an input parameter. This is a potential problem, as this number is often unknown. To overcome this problem, a number of cluster validation indices have been proposed in the literature. A cluster validation index, by definition, is a number that indicates the quality of a given clustering. Hence, if the correct number of clusters is not known, one can execute a clustering algorithm multiple times varying the number of clusters in each run from some minimum to some maximum value. For each clustering achieved under this procedure, the validation indices are computed. Eventually, the clustering that yields the best index value is returned as the final result. Cluster validation measure, such as the Dunn's index (Dunn 1973 reflects compactness, connectedness, and separation of cluster partitions.

The Dunn's index (V_D) defines the ratio between the minimum intra cluster distance to maximal inter-cluster distance, and is computed as follows:

$$(V_D) = \min_{1 \le i \le K} \left\{ \min_{1 \le j \le K, j \ne j} \left[\frac{\delta(C_i, C_j)}{\max \Delta(C_k)} \right] \right\}$$

where $\delta(C_i, C_j) = \min_{x_i \in C_i, x_j \in C_j} \lfloor d(x_i, x_j) \rfloor$

is the distance between clusters C_i and C_j (inter-cluster distance), and

$$\Delta(C_k) = \max_{x_i, x_j \in C_i} \lfloor d(x_i, x_j) \rfloor$$

is the intra-cluster distance of cluster C_k. The value of K for V_D, which is maximized, is taken as the optimal number of clusters.

2.3.3 Silhouette Score

The Silhouette Score is a measure of how much similarity an object bears to its own cluster (cohesion) compared to other clusters (separation). The values of the Silhouette Score range from -1 to +1. When the Silhouette Score is high, it indicates how well an object matches to its own cluster and how poorly it matches with the neighbouring clusters.

In our study, we calculate the Silhouette Score in the Euclidean Distance Metric.

Firstly, we compute the mean distance between $i \in C_i$ (data point i in the cluster C_i) and all other data points in the same cluster, as

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \ne j} d(i, j)$$

where $d(i, j)$ is the distance between data points i and j in the cluster C_i and $|C_i|$ indicates the number of data points in the Cluster C_i. We divide by $|C_i|-1$ as we do not include the distance $d(i,i)$ in the sum. The value $a(i)$ can be interpreted as a measure of how well i belongs to its cluster (the smaller the value, the better the belongingness).

We now compute the mean distance of point i to some cluster C_k as the mean of the distance from i to all points in C_k. That is, we compute $\frac{1}{|C_k|}\sum_{j \in C_k} d(i, j)$

For each data point $i \in C_i$, we define

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j)$$

to be the smallest mean distance of i to all points in any other cluster, and the cluster with this smallest aforementioned mean distance is said to be the neighbouring cluster of i.

The Silhouette Score of one data point i is defined as

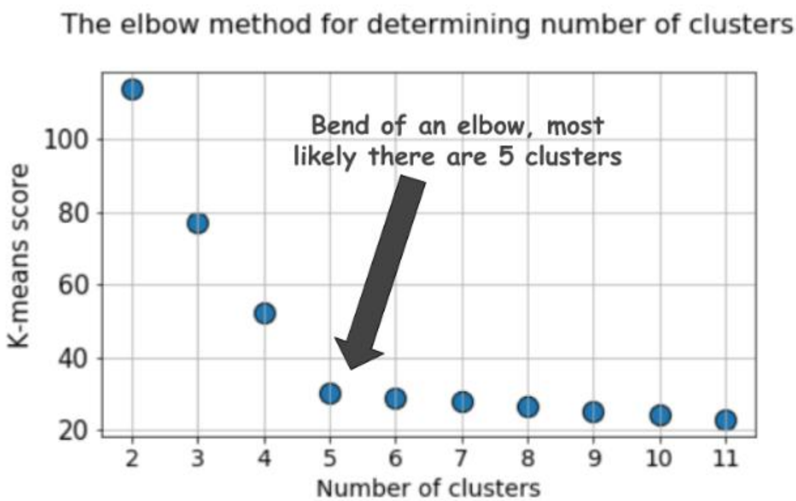
$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \text{ if } |C_1| > 1 \text{ and}$$

$$s(i) = 0 \text{, if } |C_1| = 1$$

2.3.4 The Elbow Plot

For the K-means clustering method, the most common approach for answering this question is the so-called *elbow method*. It involves running the algorithm multiple times over a loop, with an increasing number of cluster choice and then plotting a clustering score as a function of the number of clusters. The score is, in general, a measure of the input data on the K-means objective function i.e., *some form of intra-cluster distance relative to inner-cluster distance*.

Elbow method gives us an idea on what a good *k* number of clusters would be based on the sum of squared distance (SSE) between data points and their assigned clusters' centroids. We pick *k* at the spot where SSE starts to flatten out and forming an elbow.



2.4 Applications of K-Mean Clustering

It is relatively *efficient and fast*. It computes result at O(tkn), where n is number of objects or points, k is number of clusters and t is number of

iterations.

K-means clustering can be applied to machine learning or data mining

Used on acoustic data in speech understanding to convert waveforms into one of k categories (known as Vector Quantization or Image Segmentation).

Also used for choosing color palettes on old fashioned graphical display devices and Image Quantization.

K-means algorithm is useful for undirected knowledge discovery and is relatively simple. K-means has found wide spread usage in lot of fields, ranging from unsupervised learning of neural network, Pattern recognitions, Classification analysis, Artificial intelligence, image processing, machine vision, and many others.

3.1 Disadvantages of the K-Means Clustering Algorithm

The main disadvantage with the K-Means Clustering algorithm is that it is difficult to predict the K value. Furthermore, different random initializations of centroids can result in different final clusters. Also, if the original data has clusters of different size and different density, K-Means does not work well. Finally, the K-Means clustering algorithm generally provides solutions that are local optima for a given data set.

Summarily,

It takes many iterations to converge
Is very sensitive to initialization
Random initialization can easily get two centers in the same cluster
K-means gets stuck in a local optimum

3 THE K-MEANS CLUSTERING ALGORITHM APPLICABILITY CRITERION

3.1 *The Applicability Criterion*

This analysis is presented for the univariate case of dataset.

Let the data oints be represented by $x_i = 1 \, to \, n$.

Let $x_{ij-Cavg}$ be the Cluster Average of the Cluster to which x_i belongs to in the j^{th} Run of the K-Means Clustering Algorithm.

Let $x_{ij-C\max}$ be the Cluster Maximum of the Cluster to which x_i belongs to in the j^{th} Run of the K-Means Clustering Algorithm.

Let $x_{ij-C\min}$ be the Cluster Minimum of the Cluster to which x_i belongs to in the j^{th} Run of the K-Means Clustering Algorithm.

Let x_{ij-CSS} be the Cluster Silhouette Score of the Cluster to which x_i belongs to in the j^{th} Run of the K-Means Clustering Algorithm.

Let $x_{ij-CSSW}$ be the Cluster Sum Of Squares Within of the Cluster to which x_i belongs to in the j^{th} Run of the K-Means Clustering Algorithm.

We now compute the Deviations

$\delta_{ij-Cavg} = \left\{ \dfrac{\sum_j x_{ij-Cavg}}{N} \right\} - \left(x_{ij-Cavg} \right)$ where N is the number of Runs of the K-Means Clustering Algorithm.

Similarly, we compute

$$\delta_{ij-C\max} = \left\{ \frac{\sum_j x_{ij-C\max}}{N} \right\} - \left(x_{ij-C\max} \right)$$

$$\delta_{ij-C\min} = \left\{ \frac{\sum_j x_{ij-C\min}}{N} \right\} - \left(x_{ij-C\min} \right)$$

$$\delta_{ij-CSS} = \left\{ \frac{\sum_j x_{ij-CSS}}{N} \right\} - \left(x_{ij-CSS} \right)$$

$$\delta_{ij-CSSW} = \left\{ \frac{\sum_j x_{ij-CSSW}}{N} \right\} - \left(x_{ij-CSSW} \right)$$

We now Min-Max Normalize in the Range [0, 1] all the above sets of values $\delta_{ij-Cavg}$, $\delta_{ij-C\max}$, $\delta_{ij-C\min}$, δ_{ij-CSS}, $\delta_{ij-CSSW}$ separately (such aforementioned normalization done separately for each set of values). Let these thusly normalized values be represented by $\tilde{\delta}_{ij-Cavg}$, $\delta_{ij-C\max}$, $\tilde{\delta}_{ij-C\min}$, $\tilde{\delta}_{ij-CSS}$, $\tilde{\delta}_{ij-CSSW}$.

We also compute the (Sample) Standard Deviations of these aforementioned normalized sets of values, separately for each set, as:

$$\tilde{\delta}_{ij-Cavg} \rightarrow \sigma_{(ij-Cavg)S} = \left\{ \frac{\left\{ \left(\frac{\sum_j \tilde{\delta}_{ij-Cavg}}{N} \right) - \tilde{\delta}_{ij-Cavg} \right\}^2}{N-1} \right\}^{1/2}$$

$$\tilde{\delta}_{ij-C\max} \to \sigma_{(ij-C\max)S} = \left\{ \frac{\left\{ \left(\frac{\sum_j \tilde{\delta}_{ij-C\max}}{N} \right) - \tilde{\delta}_{ij-C\max} \right\}^2}{N-1} \right\}^{1/2}$$

$$\tilde{\delta}_{ij-C\min} \to \sigma_{(ij-C\min)S} = \left\{ \frac{\left\{ \left(\frac{\sum_j \tilde{\delta}_{ij-C\min}}{N} \right) - \tilde{\delta}_{ij-C\min} \right\}^2}{N-1} \right\}^{1/2}$$

$$\tilde{\delta}_{ij-CSS} \to \sigma_{(ij-CSS)S} = \left\{ \frac{\left\{ \left(\frac{\sum_j \tilde{\delta}_{ij-CSS}}{N} \right) - \tilde{\delta}_{ij-CSS} \right\}^2}{N-1} \right\}^{1/2}$$

$$\tilde{\delta}_{ij-CSSW} \to \sigma_{(ij-CSSW)S} = \left\{ \frac{\left\{ \left(\frac{\sum_j \tilde{\delta}_{ij-CSSW}}{N} \right) - \tilde{\delta}_{ij-CSSW} \right\}^2}{N-1} \right\}^{1/2}$$

As the total number of unique results possible in a K-Means Clustering Algorithm for making K Clusters with n data points is given by $m={}^{n}C_{K}$ the Population Standard Deviations of the above Sample Standard

Deviations are given by

$$\sigma_{(ij-Cavg)Pop} = \sqrt{m} \cdot \sigma_{(ij-Cavg)S}$$

$$\sigma_{(ij-C\max)Pop} = \sqrt{m} \cdot \sigma_{(ij-C\max)S}$$

$$\sigma_{(ij-C\min)Pop} = \sqrt{m} \cdot \sigma_{(ij-C\min)S}$$

$$\sigma_{(ij-CSS)Pop} = \sqrt{m} \cdot \sigma_{(ij-CSS)S}$$

$$\sigma_{(ij-CSSW)Pop} = \sqrt{m} \cdot \sigma_{(ij-CSSW)S}$$

We now Min-Max Normalize in the Range [0, 1] all the above sets of values $\sigma_{(ij-Cavg)Pop}$, $\sigma_{(ij-C\max)Pop}$, $\sigma_{(ij-C\min)Pop}$, $\sigma_{(ij-CSS)Pop}$, $\sigma_{(ij-CSSW)Pop}$ separately (such aforementioned normalization done separately for each set of values). Let these thusly normalized values be represented by $\tilde{\sigma}_{(ij-Cavg)Pop}$, $\tilde{\sigma}_{(ij-C\max)Pop}$, $\tilde{\sigma}_{(ij-C\min)Pop}$, $\tilde{\sigma}_{(ij-CSS)Pop}$, $\tilde{\sigma}_{(ij-CSSW)Pop}$

We now find a Weighted Average of $\tilde{\sigma}_{(ij-Cavg)Pop}$, $\tilde{\sigma}_{(ij-C\max)Pop}$, $\tilde{\sigma}_{(ij-C\min)Pop}$, $\tilde{\sigma}_{(ij-CSS)Pop}$, $\tilde{\sigma}_{(ij-CSSW)Pop}$ to find the value of *Variation* of results of the K-Means Clustering Algorithm Run N Times. Let this weighed average be denoted by v. We can then say $r = (1 - v)$ as the *Coefficient Of Robustness* of the results of the K-Means Clustering Algorithm for a given data set. The advantages of this value is that if the Variation of the results is high among the N runs of the K-Means Clustering Algorithm for the given data, then the results are not acceptable for reporting. Therefore, we can use this concept in specifying a value of r for each data set for running the K-Means Clustering Algorithm, so that the results are acceptable. Furthermore, for a given *r*, we can even compute the *N*.

5 REFERENCES

[1] *Lloyd, Stuart P., "Least squares quantization in PCM". Bell Telephone Laboratories Paper.* (1957)

[2] *Lloyd, Stuart P.,* "Least squares quantization in PCM", *IEEE Transactions on Information Theory, Vol 28 No 2, pp 129–137. (1982)*

[3] *MacQueen, J. B.,* Some Methods for classification and Analysis of Multivariate Observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. pp. 281–297. (1967).*

[4] Shraddha Shukla and Naganna S., A Review On K-means Data Clustering Approach, *International Journal of Information & Computation Technology*. Vol 4, No 17, pp. 1847-1860 (2014)

[5]https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a

[6] *Kriegel, Hans-Peter; Schubert, Erich; Zimek, Arthur (2016). "The (black) art of runtime evaluation: Are we comparing algorithms or implementations?". Knowledge and Information Systems.* ***52*** *(2): 341–378. doi:10.1007/s10115-016-1004-2. ISSN 0219-1377. S2CID 40772241.*

[7] *Steinhaus, Hugo (1957). "Sur la division des corps matériels en parties". Bull. Acad. Polon. Sci. (in French).* ***4*** *(12): 801–804. MR 0090073. Zbl 0079.16403.*

[8] *Forgy, Edward W. (1965). "Cluster analysis of multivariate data: efficiency versus interpretability of classifications". Biometrics.* ***21*** *(3): 768–769. JSTOR 2528559.*

[9] *Pelleg, Dan; Moore, Andrew (1999). "Accelerating exact k -means algorithms with geometric reasoning". Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '99. San Diego, California, United States: ACM Press: 277–281. doi:10.1145/312129.312248. ISBN 9781581131437. S2CID 13907420.*

[10] *MacKay, David (2003). "Chapter 20. An Example Inference Task: Clustering" (PDF). Information Theory, Inference and Learning Algorithms. Cambridge University Press. pp. 284–292. ISBN 978-0-521-64298-9. MR 2012999.*

[11] *Hartigan, J. A.; Wong, M. A. (1979). "Algorithm AS 136: A k-Means Clustering Algorithm". Journal of the Royal Statistical Society, Series C.* ***28*** *(1): 100–108. JSTOR 2346830.*

[12] *Hamerly, Greg; Elkan, Charles (2002). "Alternatives to the k-means algorithm that find better clusterings" (PDF). Proceedings of the eleventh international conference on Information and knowledge management (CIKM).*

[13] *Celebi, M. E.; Kingravi, H. A.; Vela, P. A. (2013). "A comparative study of efficient initialization methods for the k-means clustering algorithm". Expert Systems with Applications.* ***40*** *(1): 200-210. arXiv:1209.1960. doi:10.1016/j.eswa.2012.07.021. S2CID 6954668.*

[14] *Bradley, Paul S.; Fayyad, Usama M. (1998). "Refining Initial Points for k-Means Clustering". Proceedings of the Fifteenth International Conference on Machine Learning.*

[15] *Vattani, A. (2011). "k-means requires exponentially many iterations even in the plane"(PDF). Discrete and Computational Geometry.* ***45*** *(4): 596–616. doi:10.1007/s00454-011-9340-1. S2CID 42683406.*

[16] *Arthur, David; Manthey, B.; Roeglin, H. (2009). "k-means has polynomial smoothed complexity". Proceedings of the 50th Symposium on Foundations of Computer Science (FOCS). arXiv:0904.1113.*

[17] *Aloise, D.; Deshpande, A.; Hansen, P.; Popat, P. (2009). "NP-hardness of Euclidean sum-of-squares clustering". Machine Learning.* ***75*** *(2): 245–249. doi:10.1007/s10994-009-5103-0.*

[18] *Dasgupta, S.; Freund, Y. (July 2009). "Random Projection Trees for Vector Quantization". IEEE Transactions on Information Theory.* ***55*** *(7): 3229–3242. arXiv:0805.1390. doi:10.1109/TIT.2009.2021326. S2CID 666114.*

[19] **^** *Mahajan, Meena; Nimbhorkar, Prajakta; Varadarajan, Kasturi (2009). The Planar k-Means Problem is NP-Hard. Lecture Notes in Computer Science.* ***5431****. pp. 274–285. CiteSeerX 10.1.1.331.1306. doi:10.1007/978-3-642-00202-1_24. ISBN 978-3-642-00201-4.*

[20] *Inaba, M.; Katoh, N.; Imai, H. (1994). Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering. Proceedings of 10th ACM Symposium on Computational Geometry. pp. 332–339. doi:10.1145/177424.178042.*

[21] *Manning, Christopher D.; Raghavan, Prabhakar; Schütze, Hinrich (2008). Introduction to information retrieval. New York: Cambridge University Press. ISBN 978-0521865715. OCLC 190786122.*

[22] *Arthur, David; Vassilvitskii, Sergei (2006-01-01). How Slow is the k-means Method?. Proceedings of the Twenty-second Annual Symposium on Computational Geometry. SCG '06. New York, NY, USA: ACM. pp. 144–153. doi:10.1145/1137856.1137880. ISBN 978-1595933409. S2CID 3084311.*

[23] *Bhowmick, Abhishek (2009). "A theoretical analysis of Lloyd's algorithm for k-means clustering" (PDF). Archived from the original (PDF) on 2015-12-08.*

[24] *Phillips, Steven J. (2002-01-04). "Acceleration of K-Means and Related Clustering Algorithms". In Mount, David M.; Stein, Clifford (eds.). Acceleration of k-Means and Related Clustering Algorithms. Lecture Notes in Computer Science.* ***2409****. Springer Berlin Heidelberg. pp. 166–177. doi:10.1007/3-540-45643-0_13. ISBN 978-3-540-43977-6.*

[25] *Elkan, Charles (2003). "Using the triangle inequality to accelerate k-means" (PDF). Proceedings of the Twentieth International Conference on Machine Learning (ICML).*

[26] *Hamerly, Greg. "Making k-means even faster". CiteSeerX 10.1.1.187.3017.*

[27] *Hamerly, Greg; Drake, Jonathan (2015). Accelerating Lloyd's algorithm for k-means clustering. Partitional Clustering Algorithms. pp. 41–78. doi:10.1007/978-3-319-09259-1_2. ISBN 978-3-319-09258-4.*

[28] *Kanungo, Tapas; Mount, David M.; Netanyahu, Nathan S.; Piatko, Christine D.; Silverman, Ruth; Wu, Angela Y. (2002). "An efficient k-means clustering algorithm: Analysis and implementation" (PDF). IEEE Transactions on Pattern Analysis and Machine Intelligence.* ***24*** *(7): 881–892. doi:10.1109/TPAMI.2002.1017616.*

[29] *Drake, Jonathan (2012). "Accelerated k-means with adaptive distance bounds" (PDF). The 5th NIPS Workshop on Optimization for Machine Learning, OPT2012.*

[30] *Dhillon, I. S.; Modha, D. M. (2001). "Concept decompositions for large sparse text data using clustering". Machine Learning.* ***42*** *(1): 143–175. doi:10.1023/a:1007612920971.*

[31] *Steinbach, M.; Karypis, G.; Kumar, V. (2000). ""A comparison of document clustering techniques". In". KDD Workshop on Text Mining.* ***400*** *(1): 525–526.*

[32] Pelleg, D.; & Moore, A. W. (2000, June). "X-means: Extending *k*-means with Efficient Estimation of the Number of Clusters". In *ICML*, Vol. 1

[33] *Hamerly, Greg; Elkan, Charles (2004). "Learning the k in k-means" (PDF). Advances in Neural Information Processing Systems.* ***16****: 281.*

[34] *Amorim, R. C.; Mirkin, B. (2012). "Minkowski Metric, Feature Weighting and Anomalous Cluster Initialisation in k-Means Clustering". Pattern Recognition.* ***45*** *(3): 1061–1075. doi:10.1016/j.patcog.2011.08.012.*

[35] *Amorim, R. C.; Hennig, C. (2015). "Recovering the number of clusters in data sets with noise features using feature rescaling factors". Information*

Sciences. ***324****: 126–145. arXiv:1602.06989. doi:10.1016/j.ins.2015.06.039. S2CID 315803.*

[36] *Sculley, David (2010). "Web-scale k-means clustering". Proceedings of the 19th international conference on World Wide Web. ACM. pp. 1177–1178. Retrieved 2016-12-21.*

[37] *Telgarsky, Matus. "Hartigan's Method: k-means Clustering without Voronoi" (PDF).*

[38] *Aloise, Daniel; Hansen, Pierre; Liberti, Leo (2012). "An improved column generation algorithm for minimum sum-of-squares clustering". Mathematical Programming.* ***131*** *(1–2): 195–220. doi:10.1007/s10107-010-0349-7. S2CID 17550257.*

[39] *Bagirov, A. M.; Taheri, S.; Ugon, J. (2016). "Nonsmooth DC programming approach to the minimum sum-of-squares clustering problems". Pattern Recognition.* ***53****: 12–24. doi:10.1016/j.patcog.2015.11.011.*

[40] *Fränti, Pasi (2018). "Efficiency of random swap clustering". Journal of Big Data.* ***5*** *(1): 1–21. doi:10.1186/s40537-018-0122-y.*

[41] *Hansen, P.; Mladenovic, N. (2001). "J-Means: A new local search heuristic for minimum sum of squares clustering". Pattern Recognition.* ***34*** *(2): 405–413. doi:10.1016/S0031-3203(99)00216-2.*

[42] *Krishna, K.; Murty, M. N. (1999). "Genetic k-means algorithm". IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics.* ***29*** *(3): 433–439. doi:10.1109/3477.764879. PMID 18252317.*

Comments

The Notation used in the book is self-explanatory and is usually, Chapter inclusive only.

It is also to be noted that the entire core novel and new concepts research work of this book is wholly (each and every line of this research manuscript) authored by the author himself, i.e., his primary consciousness, and the author has not invited any other life entities of any kind in the participation of authoring of this book and consequently does not accord the consent of ownership of his intellectual property to any such claimants.

ABOUT THE AUTHOR

After serving in the Engineering domain as a Graduate Teaching Assistant, Assistant Professor, Associate Professor & Head during 1999-2019, now, Mr. Ramesh Chandra Bagadi is currently splitting his time as a *Technology Entrepreneur* developing solutions to latest cutting edge technologies of the futuristic kind and as an *Author Of Self Help Books.* He has received his Bachelors of Civil Engineering from Osmania University, India and a Masters each in Engineering Mechanics, Civil & Environmental Engineering, Physics from the University Of Wisconsin-Madison, USA. He is also a Registered Chartered Engineer and Fellow of The Institution Of Engineers, India.

Notes